Applicant : Boris Kalinichenko et al.   Attorney's Docket No.: 08575-103001
Serial No. : 10/809,625
Filed   : March 25, 2004
Page   : 11 of 18

REMARKS

SECTION 103 REJECTION OF CLAIM 1

Nowitz[1] draws attention to a difficulty associated with searching for media files on the internet.[2] This difficulty arises because search engines rely primarily on text.[3] *Nowitz*'s innovation lies in its "mechanism for defining a directory of a structured data store and descriptive data contained therein."[4]

Nowitz is directed to the technical problem of searching for particular content on the internet. It is thus irrelevant to Applicant's disclosure, which describes ways to make sure that content is displayed correctly on various kinds of mobile devices.

We now turn to the specific limitations of claim 1 and the textual passages from *Nowitz* that allegedly teach those limitations.

Nowitz fails to teach emulating wireless devices

Claim 1 recites the limitation of

analyzing the first set of content files for errors by emulating a first category of wireless devices;

The Examiner asserts that *Nowitz* teaches this claim limitation at col. 6, lines 3-15; col. 10, lines 18-32; and col. 11, lines 21-35, with the exception that *Nowitz* fails to teach using a wireless device.

In fact, the cited text does not describe emulating *anything*, much less "a first category of wireless devices."

The first cited passage that allegedly teaches the foregoing limitation reads as follows:

Typically, when harvesting a network for links to or metadata describing content, a web crawler or spider, crawls (searches) a network in a quasi-random fashion, following each web link it encounters. The spider uses the web link as a seed to search and retrieve information from a located resource (as a media file or

---

[1] *Nowtiz*, et al. U.S. Patent No. 7,308,464.
[2] *Nowitz*, col. 1, lines 15-20.
[3] *Nowitz*, col. 1, lines 57-64.
[4] *Nowitz*, col. 3, lines 30-33.

Applicant : Boris Kalinichenko et al.   Attorney's Docket No.: 08575-103001
Serial No. : 10/809,625
Filed   : March 25, 2004
Page   : 12 of 18

data store of media files). This process is repeated with each "seed" or link to a resource. In the workflow of the crawling process, the results discovered by the web crawler are eventually parsed and submitted to an extraction agent that further extracts metadata from the results discovered during the course of a web crawl. More information about the specifics of the workflow of a search and retrieval system is described in FIG. 6.[5]

The cited passage merely describes what a web crawler does when it visits a site. According to the passage, a web crawler follows certain links and passes certain discovered results to an extraction agent.

In the interest of clarifying the record and expediting prosecution, Applicant requests that the Examiner quote verbatim the text from the foregoing passage that allegedly teaches analyzing content files for errors, and doing so by carrying out some kind of emulation.

The second cited passage that allegedly teaches the foregoing limitation reads as follows:

The parsed results relating to the media are passed to extraction agent **668** via an extraction queue **667**. Results not associated with the media are not pursued. The extraction queue **667** comprises URIs to be analyzed with respect to associated media metadata. The entries in the extraction queue **667** are a representation of a stream and all it's associated metadata, which may comprise media URIs, Web page URIs, Web page titles, Web page keywords, Web page descriptions, media title, media author, and media genre. Each queue entry added to the extraction queue is assigned a processing time and a priority. In an exemplary embodiment of the invention, each queue entry is given a processing time of "now" and the same default priority. The iterative seeding process increases the number of queue entries added to the extraction queue **667**.[6]

This passage merely describes what happens to results that have been parsed in an effort to locate links to web resources associated with a media file. In particular, the passage describes how parsed results are placed on a queue **667** so that an extraction agent **668** can later process them.

In the interest of clarifying the record and expediting prosecution, Applicant requests that the Examiner quote verbatim the text from the foregoing passage that allegedly teaches analyzing content files for errors, and doing so by carrying out some kind of emulation.

The third passage that allegedly teaches the foregoing claim limitation reads as follows:

---

[5] *Nowitz*, col. 6, lines 3-15.
[6] *Nowitz*, col. 10, lines 18-32.

Annotation agent **678** captures and aggregates media specific metadata pertaining to the media (including multimedia and streaming media) from sources such as the media URI, the referring Web page URI, title, key words, description, and third party databases. Metadata, which may be inaccurate or "noisy," is extracted, parsed and indexed, compared with fields in known databases, and replaced with accurate metadata obtained from a valid (ground truth) database. Metadata is indexed and parsed into metadata fields and compared to other databases, such as a music or movie based database, whose accuracy is known (ground truth databases). Ground truth databases may be obtained from sources such as third party databases, previously created databases, and user entered databases, for example. Noisy fields are corrected and/or replaced with accurate data. New fields are added if appropriate.[7]

This third cited passage describes how an annotation agent **678** identifies inaccurate metadata and replaces it by replacing it with valid metadata from a database, called a "ground truth database," that is known to be accurate. Presumably, in this passage, the metadata is regarded as "the first set of content files" recited in the claim, and the identification of errors in the metadata is regarded as inherent.

Applicant points out that in the foregoing passage, the manner in which errors in metadata are identified does not require emulation. Instead, metadata errors are identified by comparing a metadata field with a corresponding field in a ground truth database. The simple act of comparing two data items hardly amounts to "emulating" anything.

In summary, none of the passages cited as allegedly teaching the limitation of "analyzing the first set of content files for errors by emulating a first category of wireless devices" actually teach anything that includes emulation.

Langer[8] teaches that wireless devices are known and that they are used for web browsing. But this teaching fails to remedy *Nowitz*'s failure to teach identifying errors through emulation. Moreover, since *Langer* describes the use of actual wireless devices, there would be no need to emulate a wireless device to see how content appears on a wireless device.

Nowitz fails to teach generating a log file
Claim 1 also recites the limitation of "generating a log file including a navigation history and error information."

---

[7] *Nowitz*, col. 11, lines 21-35
[8] *Langer*, US Patent Publication 2004/0210828.

The Examiner asserts that *Nowitz* teaches "generating a log file" at col. 11, lines 5-20 and also at col. 12, lines 14-28.

The first passage that allegedly teaches "generating a log file" reads as follows:

> If the extractor **668** is unable to open the stream, the stream is said to be invalid. If determined to be invalid, the Internet stream object is assigned a later time and priority. The Internet stream object is requeued to the extractor, and the extraction process is repeated with at the later time and in accordance with the newly assigned priority. If extraction is valid, the Internet streaming object is queued and available to all agents.
>
> A parallel part of the process determines if the accessible media file and the associated metadata links are valid. Validation comprises determining if the Web page comprises a link to a desired media file, and also determining if the desired media file works. In an exemplary embodiment of the invention, a streaming media file is retrieved and played to determine it is valid. Validation is performed periodically on all streams.[9]

The foregoing passage merely describes what happens if, for some reason, a data stream cannot be opened. According to the cited passage, the stream is placed back in the queue. The passage further describes determining whether there exists a link to a media file and validating a media file by playing that file. Applicant is unable to identify anything in the foregoing passage that even remotely resembles claim 1's limitation of "generating a log file."

In the interest of clarifying the record and expediting prosecution, Applicant requests that the Examiner quote verbatim the particular text from col. 11, lines 5-20 that is believed to disclose "generating a log file."

The second passage that allegedly teaches "generating a log file" reads as follows:

> Not all databases queried are determined to be ground truth databases. Noisy metadata (metadata that needs to be cleaned up before being processed) are compared to potential ground truth databases to determine if a potential ground truth database qualifies as a ground truth database. Noisy metadata in each field (e.g., artist, title, album) is separated into keywords by transforming any connecting characters into white space. For example, "oops_i_did_it_again" is transformed to the cleaned up "oops i did it again". Connecting characters may include, for example, period ("."), underscore ("_"), backslash ("\"), forward slash ("/"), comma (","), asterisk ("*"), hyphen ("-"), and/or any other appropriate connecting character. The fields in the noisy metadata are used to perform a full-text query against one or more fields in the potential ground

---

[9] *Nowitz*, col. 11, lines 5-20.

truth databases.[10]

The foregoing passage describes what to do with noisy metadata. According to the foregoing passage, part of what one does to noisy metadata is substituting white space for certain connecting characters.

Again, Applicant fails to see what relevance substituting a white space for a comma or asterisk might have to "generating a log file." Applicant is further unable to determine what possible relevance noisy metadata may have to "generating a log file."

In the interest of clarifying the record and expediting prosecution, Applicant requests that the Examiner quote verbatim the particular text from col. 12, lines 14-28 that is believed to disclose claim 1's limitation of "generating a log file."

Nowitz fails to teach generating a log file with navigation history and error information

Moreover, claim 1 requires not just generating any log file but generating a log file "including a navigation history and error information."

The Examiner appears to suggest that *Nowitz* teaches including a navigation history and error information at col. 6, lines 34-67. This passage, which allegedly teaches including both a navigation history and error information into a log file, reads as follows:

> An example of a structured data store in FIG. 2. FIG. 2 is an illustration of the directory structure of the www.cnn.com (CNN) web site. The structured data store of FIG. 2 is formatted into four levels. A structured data store may comprise any number of levels. Each level of a data store may comprise any number of links, objects, metadata, miscellaneous text, or any combination thereof, related to common content. An object is an entity available via the network. Examples of objects include multimedia files and streaming media files. A level may represent a web page, another web site, an object (e.g., multimedia, streaming media), metadata, miscellaneous text, or any combination thereof, encountered while conducting a crawl on a particular web site. For example, as shown in FIG. 2, the first level represents the home page **212** of the web site. Home page **212** comprises information such as the URI of the home page of the web site (i.e., www.cnn.com). The second level represents the next web page, or pages, that are accessible from the first level at that web site. Information at the second level is categorized in accordance with the type of media located at the web site. In FIG. 2, the second level comprises the media categories of video and audio. The third level represents the next web page or pages that are accessible from the second level at that web site. At the third level, information from the second level is further categorized into subcategories. As shown in FIG. 2, video media comprises the subcategories of showbiz and weather. The category of audio

---

[10] *Nowitz*, col. 12, lines 14-28.

media comprises the subcategory of showbiz. The fourth level represents the next web page or pages that are accessible from the third level at that web site. At the fourth level, each subcategory of level three is further categorized by date and subject. As shown in FIG. 2, individual media files (objects) are categorized by their respective dates (in MMDDYYYY format), and by their respect subject.[11]

It is apparent that the foregoing passage merely describes, in great detail, how the CNN web site is organized. It is unclear how one of ordinary skill in the art could learn, from the structure of the CNN web site, anything about "including a navigation history and error information" in a log file. In fact, it is unclear how a detailed discussion of CNN web site structure would suggest to one of ordinary skill in the art anything about log files in the first place.

In the interest of clarifying the record and expediting prosecution, Applicant requests that the Examiner quote verbatim that portion of the foregoing description of the CNN web site that allegedly teaches claim 1's limitation of "including a navigation history and error information into a log file."

Motivation to Combine References

One of ordinary skill in the art who was interested in ensuring that web pages displayed correctly on different kind of telephones would have had no interest in either *Nowitz* or *Langer*, much less in combining them.

From *Langer*, one of ordinary skill in the art would have learned about a web-interaction system suitable for enabling low-bandwidth devices, such as mobile telephones, to more efficiently interact with web sites. In particular, one of ordinary skill in the art would have been introduced to the idea of using a query engine that operates on XML data.

However, one of ordinary skill in the art would also have promptly recognized that *Langer*'s concern was how quickly a mobile phone could interact with a web page, *not* what the web page looked like on the mobile phone's screen. In fact, nothing in *Langer* would have even

---

[11] *Nowitz*, col. 6, lines 34-67.

alerted one of ordinary skill in the art to the possibility that a web page might display incorrectly on certain types of mobile devices.

Like *Langer*, *Nowitz* has nothing to do with the technical problem with which Applicant was concerned.

From *Nowitz*, one of ordinary skill in the art would have learned that search engines work best with textual data, and that search engines often face difficulty searching for media content. One of ordinary skill in the art would also have recognized that *Nowitz* is directed to the problem of finding content, not to making sure that the content, once found, displays correctly on different devices.

Since *Nowitz* teaches a way to search for non-textual content, and *Langer* teaches a way for a wireless phone to interact more efficiently with web pages, the result of combining the two would be a way for wireless phones to find non-textual content and to then efficiently interact with it. This has nothing whatsoever to do with the claimed invention, which is a method for *validating* wireless content.

Dependent Claims
Claims 2-18 include the limitations of claim 1 and are patentable for reasons similar to those set forth in connection with claim 1.

SECTION 103 REJECTION OF CLAIMS 19-32
Independent claims 19, 25, and 29 include limitations similar to those of claim 1 and are patentable for at least the same reasons. Claims 20-24, 26-28, and 30-32 include the limitations of their respective parent claims and are patentable for at least the same reasons.

SECTION 101 REJECTION
Applicant amends claims 19-24 as suggested by the Examiner to overcome the section 101 rejection.

SUMMARY

Now pending in this application are claims 1-32, of which claims 1, 19, 25, and 29 are independent.

No additional fees are believed to be due in connection with the filing of this response. However, to the extent fees are due, of if a refund is forthcoming, please adjust our deposit account 06-1050, referencing attorney docket "08575-103001."

Respectfully submitted,

Date: _March 21, 2008_

Faustino A. Lichauco
Reg. No. 41,942

Fish & Richardson P.C.
225 Franklin Street
Boston, MA 02110
Telephone: (617) 542-5070
Facsimile: (617) 542-8906

21882421.doc